# Memory, Attention, and Choice

Pedro Bordalo, Nicola Gennaioli, Andrei Shleifer[1]

March 9, 2017

## Abstract

We present a theory in which the choice set cues a consumer to recall a norm, and surprise relative to the norm shapes his attention and choice. We model memory based on Kahana (2012), where past experiences that are more recent or more similar to the cue are recalled and crowd out others. We model surprise relative to the norm using our salience model of attention and choice. The model predicts unstable and inconsistent behavior in new contexts, because these are evaluated relative to past norms. Under some conditions, repeated experience causes norms to adapt, inducing stable – sometimes rational –behavior across different contexts. We test some of the model's predictions using an expanded data set on rental decisions of movers between US cities first analyzed by Simonsohn and Loewenstein (2006).

---

[1] The authors are from University of Oxford, Universita Bocconi, and Harvard University, respectively. The present paper replaces, and completely reworks, a 2015 manuscript with the same title. We are grateful to Dan Benjamin, Paulo Costa, Ben Enke, Matt Gentzkow, Sam Gershman, Michael Kahana, George Loewenstein, Sendhil Mullainathan, Josh Schwartzstein, Jesse Shapiro, Jann Spiess, and Linh To for help with the paper. Shleifer thanks the Sloan Foundation and the Pershing Square Venture Fund for Research on the Foundations of Human Behavior for financial support.

## 1.Introduction

A hypothetical traveler goes to an airport for the first time, and sees expensive bottled water for sale. In standard theory, the traveler would buy the water if he is thirsty enough to be willing to pay the airport price. Introspection however suggests that the traveler might be shocked by the high price at the airport – which is much higher than the low “normal” price he is used to in stores – and not buy the water even if very thirsty. Yet the same traveler would, over time, come to find high airport prices to be “normal” (and be willing to buy water there), while also thinking that low prices are normal at stores.

A person moves from San Francisco to Pittsburgh, where rents are sharply lower. In standard theory, the mover’s reservation rent in Pittsburgh does not depend on the rent he paid in San Francisco. This mover, however, is pleasantly surprised by the great deals in Pittsburgh compared to the high “normal” level he remembers from San Francisco, and decides to rent a more expensive apartment than he would otherwise. Yet, over time, the same mover would come to find the low Pittsburgh rents “normal”, and eventually move to a cheaper apartment. Simonsohn and Loewenstein (SL 2006) offer evidence of precisely such behavior among movers between US cities.

These examples suggest that behavior may be influenced by two central mechanisms of perception and judgment: formation of norms and surprise relative to norms. These mechanisms are analyzed by Kahneman and Miller (1986) in “Norm Theory.” In their view, norms come from memory: “each event brings its own frame of reference into being” by “acting as a reminder of similar events in the past”. In turn, surprise is the cognitive reaction to an event that is very different from the norm it evokes, such as high water prices at the airport or low rents in Pittsburgh.

But how exactly are norms formed, and what events do they render surprising? What are the implications of surprise for observable economic behavior? In this paper, we address these questions by combining a textbook psychological model of memory from Kahana (2012) with Salience Theory of attention and choice from Bordalo, Gennaioli, and Shleifer (BGS 2012, 2013). Salience Theory holds that, when evaluating a choice option, our attention focuses on features that are very different from a reference point, or norm. These “surprising features” are then overweighed in choice. Salience Theory connects Kahneman and Miller’s notion of surprise, originally applied to judgments of normality, to choice.

We describe the formation of norms by adapting Kahana’s (2012) model of associative recall. This model sees episodic memory as a database of past experiences that are spontaneously recalled in response to a cue. In our setup, the cues are goods in the choice set, each of which evokes memories of past experiences with the same good. Cued recall follows two fundamental principles. First, it is associative and driven by similarity. That is, a cue leads to recollection of similar past experiences, where similarity is defined along hedonic attributes such as the quality and price of a good, as well as along contextual attributes such as location and time of the experience. Second, recall is subject to interference. This means that experiences more similar to the cue block recall of other, less similar, experiences. This model yields several basic features of recall, such as recency effects and positive effects of repetition on subsequent recall. The memories evoked by the stimulus are then aggregated into a norm, in line with Kahneman and Miller’s idea that norms are built by “selectively evoking stored representations” of similar past experiences.

The fact that recall of past experiences is driven by their similarity to the cue has profound implications. Because similarity may be strong along contextual attributes such as time or location, recall can bring to mind norms that have very different hedonic attributes such as price or quality than the current choice options. Surprising qualities or prices then distort the weight the consumer attaches to these attributes and thereby his choice. In contrast, when similarity along hedonic attributes such as quality and price dominates recall, the current options simply bring to mind their own past occurrences. In this case, we say that the consumer is well adapted to the choice environment.

Similarity-based recall then implies that personal history frames the consumer’s choice. When facing an entirely new choice setting, the consumer’s norm – coming from memory – is not adapted. This is the case with the first-time traveler at the airport, for whom only downtown prices are available in memory. Similarity here acts through recency, generating a “low price” norm for water. The consumer gasps at the high airport price, overweighs it in choice, and chooses not to buy. This memory-based mechanism for surprise accounts for Simonson and Tversky's (1993) background context effects, where recall of previous choices shapes the evaluation of current ones.

Because similarity depends on hedonic attributes, repeated and frequent experience with different situations also leads to the adaptation of norms. For a traveler frequently visiting airports, seeing a high price mainly reminds him of past airport experiences. Likewise, frequent experiences at stores causes low prices to cue previously experienced low prices. Under some conditions, norms fully adapt (ex post) to the reality the consumer faces. This has two implications. First, consumer behavior looks more rational in that it is stable in each location, independent of the most recent experience. Second, the higher price norm at the airport implies that the traveler will eventually be willing to pay more for the same water there than at the store, even if equally thirsty. This memory-based mechanism accounts for Thaler's (1985) famous experiment, in which willingness to pay for beer on the beach depends on reminders of where the beer is bought (resort versus store), rather than reflecting a stable reservation price.

We take the predictions of our model to the data, revisiting the Simonsohn and Loewenstein (SL 2006) evidence on rental expenditures of movers between US cities from the Panel Study of Income Dynamics (PSID), with twenty additional years of data. SL find that movers to a new city choose housing with rents that are closer to those in their city of origin, relative to other households with similar incomes and demographic characteristics. On their subsequent moves within the city, however, the rental choices of movers are no longer shaped by past prices. Both findings are predicted by our model, and we confirm both in more recent data. We also test two novel predictions of situation-specific norms. First, movers are better adapted to their new location, and thus past rents are less important, when they have previously lived in a city with similar rents. Second, price in the city of origin has a stronger effect on rent paid for movers to cheap cities than for movers to more expensive ones. There is some support for both implications in the data.

In our model, price and quality norms effectively act as reference points that shape the decision maker's locus of attention. In this sense, our memory-based approach provides a unified way to think about different types of reference points proposed in the vast literature on this topic. The "status quo" view of reference points adopted in Prospect Theory (KT 1979) corresponds to the case in which a stable history of past experiences fully determines the memory-based reference. This “backward looking” view is supported by a substantial empirical literature (e.g., Genesove and Mayer 2001). However, even backward looking reference points eventually adapt to new settings (e.g. DellaVigna et al 2017). In our model, both “status quo” and “slow adaptation” effects are driven by similarity in recall: similar prices and qualities crowd out dissimilar ones, provided they are recent enough. But our model also predicts how and when reference points become situation-specific, and independent of recent experiences. This ex-post adaptation is not easily reconcilable with the standard backward looking view, and generates consistency of references – and stability of choice – in a well-defined set of circumstances.

Koszegi and Rabin (2006) develop a purely forward-looking approach to reference points based on rational expectations, in part to account for the prevalence of well-calibrated reference points. Although expectations clearly play a role in reference point formation, this approach is difficult to reconcile with the evidence that normatively irrelevant backward looking anchors influence choice. In our model, reference points are shaped by both recency and choice similarity. This approach provides a middle ground between mechanical adaptation and rational expectations. In some cases, the memory based reference point is influenced by historical anchors; in others it is well adapted, resembling rational expectations. The properties of associative memory yield predictions for when norms should or should not be fully adapted to the choice set, depending on the frequency with which the decision maker is exposed to different choice sets.

In the next section, we summarize the research on memory that motivates our formulation, focusing on Kahana’s (2012) model. Also in that section, we apply this model to build a theory of norms, and develop some of the key predictions. In Section 3, we describe how norms generate reference points, and combine this theory with salience theory to formulate a model of choice. In Section 4, we apply this model to the Simonsohn and Loewenstein analysis of movers between US cities. Section 5 concludes.

## 2. The Model

### 2.1 Similarity-Based Recall in Psychology

Since the 1880s, a large body of experimental work has described the workings of episodic memory, or memory of past experiences. Recently, this evidence has led to the development of formal models of recall based on similarity. In these models, memory is viewed as a storage-and-retrieval facility for experiences/events. Events are encoded as "memory traces", which are vectors of attributes. Some attributes are inherent to the event, others are contextual. For instance, when drinking a glass of milk our memory records the taste and color of the milk (intrinsic attributes), but also contextual conditions such as day and location. The memory database can thus be described as an *event x attribute* matrix.[2]

Recall is a spontaneous and subconscious process in which a current experience stimulates the retrieval of a trace from memory. As described by Kahana (2012), recall obeys two principles:

1. Recall is associative, driven by similarity: presenting a stimulus facilitates recall of items from memory that are similar to that stimulus.
2. Recall is subject to interference: recall of items of given similarity to the stimulus is weakened or blocked entirely in the presence of more similar items in working memory.

To illustrate, consider the three prominent experimental paradigms used to study memory. In item recognition tests, subjects assess whether given words were part of previously shown lists of words. These tests illustrate the role of similarity because i) the probability of recall is higher for items that belong to the list, and ii) subjects are more likely to mistakenly recognize words if these are similar to a list member (e.g. they recognize yoghurt when milk is on the list). In cued recall, subjects retrieve words that are pairwise associated with a cue, having previously been shown lists of relevant word pairs. These tests illustrate the role of interference: when a certain cue $A$ becomes associated with more and more words, recall of old

---

[2] Similarity models focus on episodic memory (i.e. memories of past experiences), while leaving out so-called semantic memory, a broad term that covers functional associations and rule based thinking (e.g., recalling "glass" after seeing "milk"). Semantic memory allows humans to create mental models, which may play a role in reference point formation. However, it raises difficult theoretical issues, and is perhaps orthogonal to the mechanisms of experience-based reference points that we focus on here.

associations declines (the fan effect). These interference effects are stronger for items that are more similar, so similarity shapes cued recall as well. Finally, in free recall tests, subjects retrieve as many words as possible from a long list. Here, previously recalled words act as cues for further recall, and the observed sequences of recalled words are well accounted for by the forces of similarity.

What are the defining features of similarity? In the standard model, similarity decreases with distance in the space of intrinsic *and* contextual attributes. A key feature of associative memory is that recall can spontaneously bring to mind experiences that are similar along intrinsic attributes (and potentially relevant for the current choice), but also experiences similar along contextual, and normatively irrelevant, attributes. Contextual attributes refer to features of the environment in which the item is presented. These can be an individual's mood, or information about the physical environment (e.g. location, weather, or time of day). The "contextual drift hypothesis" views the time of the experience as a key context attribute because both internal and external aspects of context change slowly over time. This creates recency effects: similarity to the current experience is higher for more recent events, so their recall is more likely.[3]

The resulting similarity model parsimoniously accounts for the large body of experimental evidence, as well as for the two most basic observations about memory: the laws of repetition and of recency. The recency of an experience augments its similarity to any current cue. Furthermore, repetition of an event creates many similar traces in memory, thus enhancing their ability to crowd out other traces.

A key contribution of our paper is to incorporate a standard model of memory from psychology into an economic decision-making setting. Only a few economic models have previously explicitly dealt with memory. Early papers incorporating memory limitations explored optimal storage of information given limited capacity (e.g., Dow 1991, Wilson 2014) or analyzed decision problems with exogenous imperfect recall (Piccione and Rubinstein 1997). Rubinstein (1998) summarizes some of this literature. In case based decision theory (CBDT, Gilboa Schmeidler 1995), decision makers recall past experiences based on their

---

[3] This approach has challenged the traditional view that recency effects stem from memory decay, i.e. forgetting (see also Brown, Chater, and Neath, 2007), and replaces it with interference from more recent experiences.

similarity with the current problem, but similarity is characterized axiomatically, not psychologically. For example, CBDT does not allow for contextual attributes to influence recall.

A more recent literature takes a more psychological approach to memory. In Mullainathan (2002), limited memory distorts Bayesian updating and forecasting of an economic variable. Mullainathan's model allows for similarity to influence recall, but his notion of similarity includes neither context nor interference. Taubinsky (2014) studies optimal reminders in a model where memory is imperfect and mental rehearsal promotes recall. Ericson (2016) studies the interaction of forgetting and procrastination, drawing implications for the demand for reminders. These models abstract from both similarity and interference. The marketing literature also discusses the role of past purchases in current reference price, but the models of memory used there are centrally focused on recency effects (for a review, see Mazumdar et al 2005). Finally, Malmendier and Nagel (2011, 2016) document that expectations of stock market performance and of inflation are disproportionately shaped by past experiences of those variables, suggesting that own experience plays an outsized role in shaping beliefs.

A number of recent papers build on Kahneman and Tversky's (1972) representativeness heuristic to explore how selective memory shapes beliefs. In this approach, representativeness distorts beliefs by highlighting the features that are most diagnostic of, or similar to, a group in contrast to a comparison group (Gennaioli and Shleifer 2010, Bordalo, Coffman, Gennaioli, Shleifer 2016). While such effects could also be at play in the recall of norms, our approach, like Norm Theory, abstracts from them.

### 2.2 Memory Database

In line with Kahana's model of associative recall, memory is a database of past experiences, which we restrict to past choice situations. Observing a certain choice option $j$ (e.g., a bottle of beer of brand $j$) at time $t$ characterized by quality and price attributes $(q_{jt}, p_{jt})$ leaves a "memory trace" $(q_{jt}, p_{jt}, \boldsymbol{c})$ in the agent's database. In this notation, $\boldsymbol{c}$ is a "context" vector capturing non-hedonic attributes present during encoding. This vector could parsimoniously be defined by the time $t$ of the experience and its location $s$, $\boldsymbol{c} =$

$(t, s)$, where location captures the "type of store" (e.g., convenience versus airport) or "city" (e.g., San Francisco versus Pittsburgh) where the good was observed. Location is a relevant context factor in some of our applications, but it does not alter the predictions based on the time factor alone. To streamline our analysis, we make the minimal assumption that context is only defined by the time of the experience, and write $\boldsymbol{c} = t$. We later return to the role of location as context.

Denote by $T_j = \{t_1, t_2, \dots, t_n\}$ the ordered set of all dates at which good $j$ was observed in the past. A generic memory trace consists of a triplet of real numbers, $q_{jt_r}, p_{jt_r}, t_r \in \mathbb{R}$, where $p_{jt_r}$ and $q_{jt_r}$ denote the price and quality of the $r$th observation of the good, which occurred at time $t_r \in T_j$. After observing the good $n$ times, the memory database for good $j$ at $t > t_n$ is the matrix:

$$M_{jt} = \begin{pmatrix} p_{jt_1} & p_{jt_2} & & p_{jt_n} \\ q_{jt_1} & q_{jt_2} & \dots & q_{jt_n} \\ t_1 & t_2 & & t_n \end{pmatrix},$$

which lists all past experiences of good $j$. Database $M_t$ adds past experiences of all goods $M_t \equiv \{M_{jt}\}_{j \in J}$.[4]

To illustrate, consider a consumer who visited a convenience store $n$ times in the past. At the store, he considered a bottle of water of constant quality and price $q, p > 0$. Thus, the consumer's memory database of good $j =$ water bottles at time $t$ consists of repeated experiences of a single good:

$$M_{jt} = \begin{pmatrix} p & & p \\ q & \dots & q \\ t_1 & & t_n \end{pmatrix}, \tag{1}$$

Suppose now that, at time $t > t_n$, this consumer has an $n + 1$th experience in an airport in which the price of the same water bottle is marked up to $p' = p + \Delta$. This experience is then encoded in the memory database which, for $t' > t$ (and prior to the next experience) is given by:

---

[4] In principle, the assignment of experiences into categories of goods is a function of experience itself. For example, a person trying wine for the first time may classify it as a "drink", but will at some point create a "wine" category, and eventually "white wine" category, and so on. Here, we take this categorization as given, reflecting our interest in choice among familiar goods, while varying choice contexts, as illustrated in the water example.

$$M_{jt'} = \begin{pmatrix} p & p & p+\Delta \\ q & \dots q & q \\ t_1 & t_n & t \end{pmatrix}. \tag{2}$$

As the consumer visits more airports, he keeps adding "airport water" vectors to his database regardless whether water is bought or not. For the purposes of recall, the "experience of observing a good" is broader than the mere act of buying it. Considering the good for choice, seeing its price in an advertising campaign, or being told by a friend about it, all leave memory traces that may be recalled when shopping. Other kinds of experiences may also leave traces in memory: rehearsal of past experiences, promises, or future goals. These can also be included in the model, albeit with some modifications. To focus only on the consumer's observable history, we restrict $M_t$ to comprise past choices involving good $j$.

### 2.3 Cued Recall

Observing a stimulus cues spontaneous recall of past experiences similar to that stimulus. In our setting, stimuli are goods in a choice set $\boldsymbol{C}$. Because recall operates at the level of a single stimulus, in this section we simplify notation and omit the good's index $j$, with the understanding that all experiences refer to the same good. (We reinstate this index in section 3 when we consider choice among several options.) Denote by $e_t = (q_t, p_t, t)$ the experience of seeing good $j \in \boldsymbol{C}$ at time $t$. This experience sets off a recall process from that good's memory database $M_t$ that follows two principles: similarity and interference.[5] We formalize similarity as follows.

**Definition 1** (Similarity) *The similarity of past experience* $e_{t_r} \equiv (q_{t_r}, p_{t_r}, t_r)$ *in* $M_t$ *to the current experience* $e_t \equiv (q_t, p_t, t)$ *of the same good j is measured by*

$$S(e_{t_r}, e_t) \equiv S(|q_t - q_{t_r}|, |p_t - p_{t_r}|, |t - t_r|), \tag{3}$$

*where the function* $S: \mathbb{R}^3_+ \rightarrow [0,1]$ *decreases in each of its arguments, and* $S(0,0,0) = 1$.

---

[5] The assumption that observing good $j$ cues recall of the same good is a reduced form of capturing semantic memory. This assumption is also made in Norm Theory, which restricts memory-based norms to elements of the same category, e.g. "the norm for horses should not include carriages."

Similarity between two experiences increases if they get closer along hedonic or contextual (as proxied by time) dimensions. The so-called "geometric approach" to similarity in (3) is standard in the psychology and neuroscience literature on memory. Kahana (2012) proposes the metric $S(e_{t_r}, e_t) = e^{-\tau d(e_{t_r}, e_t)}$, where $d(e_{t_r}, e_t)$ is the Euclidean distance between $e_{t_r}$ and $e_t$.[6] Due to similarity along contextual dimensions, recall can bring to mind experiences that differ from the cue along hedonic attributes. As stressed above, this is a key feature of associative memory.

We next describe how norms are formed by cued recall, and the role of interference in that process. The current experience $e_{jt}$ activates past experiences in the memory database $M_t$ to different degrees, depending on similarity. As in Kahneman and Miller, the norm aggregates past experiences by attaching a larger weight to those that have higher degree of activation (i.e. the more available ones). Interference refers to the phenomenon that past experiences that are similar to the stimulus $e_{jt}$ reduce the availability of less similar ones and thus play an outsized role in the norm.

**Definition 2.** *The activation of a past experience* $e_{t_r}$ in $M_t$ *by current experience* $e_t$ *is:*

$$h_{t_r} = h\left(S(e_{t_r}, e_t)\right),$$

*where* $h(.): [0,1] \to \mathbb{R}_+$ *is increasing. The recalled norm is a similarity-weighted average* $e^m(e_t)$*:*

$$e^m(e_t) = \sum\nolimits_{t_r \in T} e_{t_r} * w_{t_r} \, . \tag{4}$$

*where the weight attached to experience* $e_{t_r}$ *is its relative activation* $w_{t_r} = \frac{h_{t_r}}{\sum_{t_s \in T} h_{t_s}}$.

---

[6] Here $\tau$ is a constant that maps distance to (log) similarity. This approach follows multidimensional scaling models (Torgerson 1958). Definition 1 nests also weighted Euclidean models in which similarity decreases in the metric $\sqrt{\lambda_q (q_t - q_{t_r})^2 + \lambda_p (p_t - p_{t_r})^2 + \lambda_t (t - t_r)^2}$. The weights capture the unequal importance or salience attached to the different attributes. Tversky (1977) highlights cases in which judgments of similarity do not follow geometric properties. He proposes a contrast model in which similarity between two experiences may not be symmetric and depends on other experiences being considered at the same time. Such contextual factors could in principle be captured in the above through the weights $\lambda$.

The norm $e^m(e_t)$ evoked by the current experience $e_t$ satisfies two properties. First, the weight attached to particular past experiences increases with their similarity to the current experience. Second, by weight normalization, the weight attached to a particular past experience decreases in the similarity between other experiences and the cue $e_t$. This is the interference effect.

Definition 2 is sufficient to obtain our results. For concreteness, in the following we consider the activation function $h\left(S\left(e_{t_r}, e_t\right)\right) = S\left(e_{t_r}, e_t\right)^{\eta}$, with $\eta \geq 0$. It satisfies Definition 2, and it implies that the weight attached to good $e_{t_r}$ is given by:

$$w_{t_r} = \frac{S\left(e_{t_r}, e_j\right)^{\eta}}{\sum_{t_s \in T} S\left(e_{t_s}, e_j\right)^{\eta}},$$

In this specification, interference increases in $\eta$ in the sense that the elasticity of $w_{t_r}$ to a marginal increase in the similarity of any other experience $e_{t_u}$ is equal to $-\eta w_{t_u}$. As $\eta \rightarrow \infty$, interference becomes so strong that $e^m(e_t)$ converges to the most similar experience, namely $e^m(e_t) \rightarrow \operatorname{argmax} S\left(e_{t_r}, e_t\right)$.[7]

In Kahana's (2012) model of probabilistic recall, past experience $e_{t_r}$ is recalled with probability $\frac{S(e_{t_r}, e_t)}{\sum_{t_s \in T} S\left(e_{t_s}, e_t\right)}$, which captures relative similarity to the cue $e_t$. Our specification nests this model for $\eta = 1$, which means that the norm is the average good recalled by a subject sampling his memories.

Our approach makes some simplifying assumptions concerning the encoding of memories, and their subsequent availability for recall. It abstracts from the possibility that events that are distinct or surprising leave stronger traces in memory and thus can more easily be retrieved, as in the "peak-end" rule of recall (Kahneman et al. 1993).[8] It also abstracts from mental rehearsal about experiences driving their availability

---

[7] Definition 1 does not nest the truncation model used by Gennaioli and Shleifer (2010), in which the $K \geq 1$ most similar items are recalled and the $|T| - K$ least similar experiences are forgotten. This case features a stronger version of interference, in which the activation of $e_{jt_r}$ falls with the similarity of other memory traces to the cue. Under mild further conditions, our main results also hold under more general activation functions $h_{t_r} = h\left(S\left(e_{t_r}, e_t\right); \left\{S\left(e_{t_s}, e_t\right)\right\}_{s \neq r}\right)$ that nest the truncation model.

[8] It is indeed easier to recall surprising events, but this is probably not a key driving force of memory-based norms. A meal at an extraordinary restaurant is memorable, but it does not alter our norm for restaurant meals, which is based on recall of more ordinary restaurants. Still, this aspect can be captured in our model by assuming that activation of a past experience $e_{t_r}$ increases in the distance between $e_{t_r}$ and the norm $e^m\left(e_{t_r}\right)$ it evoked in the database $M_{t_r}$ available

(see Mullainathan 2002). Finally, it assumes each experience is a primitive, without allowing for the possibility that a decision maker may not notice – and thus not encode – certain attributes (Schwartzstein 2014), or that attributes may be encoded separately (Bushong and Gagnon-Bartsch 2016). Future work may enrich the model along these lines.

Definition 2 yields two key laws of recall, namely that it is facilitated by an experience's recency and by how often it was repeated in the past.

**Proposition 1** (Laws of Recency and Repetition). *Denote by* $(\hat{q}, \hat{p})$ *a quality price pair experienced in the past for good j, and by* $w(\hat{q}, \hat{p})$ *the total weight on all experiences of* $(\hat{q}, \hat{p})$*. Then,* $w(\hat{q}, \hat{p})$ *weakly increases if:*

*i) the* $(\hat{q}, \hat{p})$ *pair has been observed more recently.*

*ii) the* $(\hat{q}, \hat{p})$ *pair has been observed more frequently in the past.*

The law of recency holds because recent events are very similar along the time dimension. When time $t'$ at which $\hat{q}, \hat{p}$ was observed gets closer to the present, the trace $(\hat{q}, \hat{p}, t')$ becomes more similar to the current experience $(q_t, p_t, t)$ along the time dimension. This facilitates the activation of $\hat{q}, \hat{p}$, increasing its weight in the norm. The law of repetition holds because adding multiple experiences of quality-price pair $\hat{q}, \hat{p}$ to the memory database weakly increases the number of times it enters the norm and thus its total weight.

### 2.4 Norms

We have just described how experience $e_t = (q_t, p_t, t)$ evokes its norm $e^m(e_t) = \left(q_t^m, p_t^m, t_j^m\right)$ by cueing the spontaneous recall of similar experiences. In what follows, we use the expression "norm for good $(q_t, p_t)$" as the vector of hedonic attributes $(q_t^m, p_t^m)$ computed according to Definition 2:

$$q_t^m = \sum\nolimits_{t_r \in T} q_{t_r} * w_{t_r} \ , \quad p_t^m = \sum\nolimits_{t_r \in T} p_{jt_r} * w_{t_r}.$$

---

at time $t_r$. Formally, surprising experiences in which $\left|e_{t_r} - e^m\left(e_{t_r}\right)\right|$ is larger would more likely be recalled for any given subsequent experience $e_t$ (see Bushong and Gagnon-Bartsch 2016 for a related approach).

We next describe several key properties of memory-based norms.

Consider first how the features of the consumer's database shapes norms. Suppose that the choice environment is stable, in the sense that the same set of goods was previously observed repeatedly. This is the case of the first-time traveler described in Equation (1), where all experiences are of the form $(q, p)$ and the memory database is $M_t \equiv (q, p, t_r)_{r=1,\ldots,n}$. In this case, the norm for *any* current experience $e_t$ consists of the "status quo" quality and price, $(q, p)$. When choice experiences are stable, memory-based norms yield the "status quo" of Kahneman and Tversky (1979).

In a changing environment, in contrast, the norm adapts. Once a different experience has entered the database, it becomes available for recall and influences the future norm. This is not a mechanical, backward looking convergence of norms to the past. Rather, the adaptation of norms is molded by the current experience $e_t$, which cues recall of experiences that are similar to itself. Proposition 2 describes this mechanism.

**Proposition 2.** *Let $M_t$ be a memory database at t and let $\widehat{M}_t$ be a memory database at the same date obtained by adding past experience $e_{\hat{t}} = (\hat{q}, \hat{p}, \hat{t})$ to $M_t$ (namely, $\widehat{M}_t = M_t \cup \{e_{\hat{t}}\}$). For $\eta < \infty$, we have:*

*i) Relative to $M_t$, the norm under $\widehat{M}_t$ attaches a higher weight to $(\hat{q}, \hat{p})$.*

*ii) Adaptation is shaped by hedonic similarity to the cue: the weight attached to $(\hat{q}, \hat{p})$ increases in the similarity between $(\hat{q}, \hat{p})$ and the hedonic attributes $(q_t, p_t)$ of the cue.*

Proposition 2 immediately implies that an important source of adaptation is the repetition effect of Proposition 1. This is point i). The second, key source of adaptation is similarity: experience $e_{\hat{t}}$ weighs more on the current norm the stronger its similarity to the cue $e_t$. Similarity in the time dimension, which decreases in $|\hat{t} - t|$, gives rise to recency effects as in Proposition 1. Additionally, Proposition 2 highlights the role of attribute similarity: $e_{\hat{t}}$ weighs more on the norm cued by $e_t$when $|q - \hat{q}|$ and $|p - \hat{p}|$ are smaller.

To illustrate these effects, consider the adaptation of the traveler's price norm. In the first airport visit, the price norm was the status quo downtown price $p_t^m = p$. In the next shopping moment $t'$ property

i) implies that the consumer's price norm partially adapts to $p_{t'}^m = p + w\Delta$. Here, $w > 0$ is the weight put on the high airport price, which is now included in memory and available for recall. Crucially, the weight $w$ depends on the similarity between the high airport price and the current price cue (point ii). If at $t'$ the consumer visits the airport again, the price cue $p + \Delta$ is very similar to the past airport price, so $w$ is large. If at $t'$ the consumer shops downtown, the price cue $p$ is dissimilar from the past airport price, so $w$ is low. As a result, similarity triggers recall of high prices at the airport, and the recall of low prices downtown, generating situation-specific adaptation of norms.

In fact, selective adaptation may cause reference points to fully adapt to different environments.

**Corollary 1.** *Suppose that after experience $e_{\hat{t}} = (\hat{q}, \hat{p}, \hat{t})$ the consumer experiences $e_t = (\hat{q}, \hat{p}, t)$. The norm at t is then fully adapted, that is, equal to the currently observed hedonic attributes $(\hat{q}, \hat{p})$, provided:*

*i) Similarity in $(q, p)$ is stronger than recency, $S(0,0, |\hat{t} - t|) > \max_r S(|q_{t_r} - \hat{q}|, |p_{t_r} - \hat{p}|, |t_r - t|)$, where $\hat{t}$ is the date of the most recent observation of $(\hat{q}, \hat{p})$.*

*ii) Interference is maximal, $\eta \to \infty$.*

Full adaptation to a quality-price profile is a fixed point of recall whereby observing $(\hat{q}, \hat{p})$ only triggers the recall of $(\hat{q}, \hat{p})$ itself. Corollary 1 highlights the conditions under which a consumer fully and immediately adapts to such a quality-price profile, even if the profile was surprising the first time it was seen. Two conditions are required for such extreme adaptation. First, price and quality similarity must be stronger than recency, as in point i). In this case, the next time the consumer encounters $(\hat{q}, \hat{p})$, associative recall favors retrieval of the same past experience $(\hat{q}, \hat{p})$ rather than that most recently observed. Second, interference must be very strong, as in ii), so the most available memory $(\hat{q}, \hat{p})$ crowds out all the others.

The conditions of Corollary 1 are admittedly extreme, but they illustrate how selective adaptation can account for Kahneman's observation that "we are surprised only once". With similarity-based recall, each situation triggers a different filtering of the memory database, and pushes different memories to the fore of consciousness. A fully adapted consumer has low prices in mind when downtown, and high prices in mind when at the airport. More broadly, even an unlikely event that is surprising the first time may look "normal" on its second occurrence, because it triggers recall of itself.[9]

The idea that consumers may have situation-specific and well-calibrated norms about prices has motivated the rational expectations approach to reference points (Koszegi and Rabin 2006).[10] Memory based norms coincide with rational expectations when the consumer is adapted as in Corollary 1. Such adaptation is even stronger if location serves as a context attribute because then being at the airport (store) already primes the consumer to think of high (low) prices, due to location similarity.[11] Still, the stringent conditions of Corollary 1 suggest that full adaptation is rare. If adaptation is partial, important differences arise relative to the rational expectations predictions: cues trigger the spontaneous recall of past experiences which can act as normatively irrelevant anchors for valuation and choice.

### 3. Norms, Attention and Choice

In line with Norm Theory, choice is a two-step mental process: in the first step, considered in Section 2, the choice set cues recall of similar goods experienced in the past. In the second step, the evaluation of available options is shaped by how surprising their attributes are perceived to be *relative to the recalled norms*.[12] We model this second step using Salience Theory (BGS 2012, 2013), which describes how the comparison between a choice option and a reference option generates surprises, shapes valuation, and drives choice.

---

[9] Kahneman (2003) offers an auto-biographical example of this point. Having once seen a burning car on the side of a road, he half-expected to see it again when driving by the same spot (and might thus not be surprised if he did).

[10] Other models map reference points to rational expectations. Bell (1985) identifies reference price as the rational expected price (see also Gul 1991). Barberis and Huang (2001) and Barberis and Xiong (2009) take a related approach in asset pricing with respect to the expected risk free rate.

[11] With stochastic prices our model predicts that consumers can be *too well* adapted: faced with a price realization, consumers recall not the expected price but rather realizations that are similar to that observed. As a result, the consumer may adapt to each possible realization and be relatively insensitive to the expected price.

[12] While it is in principle possible to elicit memory-based norms, reference points are not directly observable. Existing work on reference points thus tests for joint hypotheses of a model of reference points and a model of reference dependent valuation (typically loss aversion). Here we follow the same strategy.

At time $t$, the consumer must choose one item from a set $\boldsymbol{C}(t) = \{(q_{jt}, p_{jt})\}_{j=1,\dots,J}$ of $J$ goods characterized by their (known) quality and price. We assume that each option $(q_{jt}, p_{jt})$ in the choice set acts as a cue, evoking the corresponding norm $(q_{jt}^m, p_{jt}^m)$ from memory. In line with BGS (2012, 2013), the consumer evaluates each good in $\boldsymbol{C}(t)$ by overweighting its most salient attribute, namely the one that stands out the most relative to the set of norms $\{(q_{jt}^m, p_{jt}^m)\}_{j=1,\dots,J}$ that is recalled.[13] Note that in this model each good is compared to all norms, not only to its own norm. This assumption captures a key feature of the psychology of attention: attention is directed to features that are salient with respect to the entire choice context, here captured by the set of recalled goods. Thus, if one good has a much lower price (yet similar quality) than another, the latter will seem expensive in comparison, shaping attention and valuation. This mechanism accounts for context effects in choice (e.g. the decoy effect), as well as Simonson and Tversky's (1993) "background context" effect in which past choice sets come to mind.

The set of norms is summarized in terms of the average norm, which we refer to as the *memory-based reference point*:

$$MBRP(\boldsymbol{C}(t)) = (\bar{q}_t^m, \bar{p}_t^m), \qquad where \quad \bar{q}_t^m = \frac{1}{J}\sum_j q_{jt}^m, \quad \bar{p}_t^m = \frac{1}{J}\sum_j p_{jt}^m$$

The memory-based reference vector $(\bar{q}_t^m, \bar{p}_t^m)$ consists of the average, or normal, levels of quality and price across all the experiences that come to mind. $(\bar{q}_t^m, \bar{p}_t^m)$ yields a ratio of quality to price that is perceived as normal in the current choice.

For each good $(q_{jt}, p_{jt})$, the salience of quality is then $\sigma(q_{jt}, \bar{q}_t^m)$ and that of price is $\sigma(p_{jt}, \bar{p}_t^m)$, where $\sigma$ is a salience function that measures the proportional distance between attributes and their reference levels.[14] Up to normalization, option $(q_{jt}, p_{jt})$ is evaluated as:

---

[13] We depart from BGS (2013), which assumes that salience is defined with respect to the centroid of the union of the choice set and the set of norms (which we called the evoked set). Our present specification, in which the reference only consists of the centroid of norms, simplifies the analysis without changing our main results. In settings where choice set effects matter, such as decoy effects, the original definition is necessary.

[14] Formally $\sigma(x, y)$ is symmetric, homogeneous of degree zero, and increasing in the ratio $x/y$ for $x \geq y$. These properties imply that salience displays ordering, $\sigma(x, y) > \sigma(x', y')$ for any $x > x' > y' > y \geq 0$, and diminishing

$$\sigma\left(q_{jt}, \bar{q}_t^m\right) \cdot q_{jt} - \sigma\left(p_{jt}, \bar{p}_t^m\right) \cdot p_{jt}. \tag{5}$$

The consumer chooses the good in the choice set $\boldsymbol{C}(t)$ that maximizes (5). If price and quality are equally salient for good $j$, $\sigma\left(q_{jt}, \bar{q}_t^m\right) = \sigma\left(p_{jt}, \bar{p}_t^m\right)$, the consumer's valuation of good $j$ is proportional to the rational one, so it rationally trades off the good's quality and price. If quality is more salient than price, $\sigma\left(q_{jt}, \bar{q}_t^m\right) > \sigma\left(p_{jt}, \bar{p}_t^m\right)$, the consumer overweighs quality, and conversely if price is salient.

Equation (5) illustrates how memory shapes valuation and choice. Selective recall, which is a function both of the choice set and of the consumer's history, determines the reference quality $\bar{q}_t^m$ and the reference price $\bar{p}_t^m$. These memory-based references then distort valuation: for each option, disproportionate attention is paid to the attribute that is most different, or salient, *relative to the reference*.[15]

To build intuition, consider the valuation of a good that presents a quality-price trade-off relative to the reference $(\bar{q}_t^m, \bar{p}_t^m)$. The homogeneity of degree zero of the salience function then implies that the advantage of good $j$ over the reference, higher quality or lower price, is salient provided

$$\frac{q_{jt}}{p_{jt}} > \frac{\bar{q}_t^m}{\bar{p}_t^m},$$

namely, when good $j$ has a higher quality to price ratio then the average good recalled from memory. Intuitively, in this case the good is perceived as a good deal, providing more quality per unit cost.

The model generates memory-based context effects. For example, a high quality and expensive option may look like a better deal to a consumer who has previously experienced relatively high prices. Recall of such prices inflates the price norm $\bar{p}_t^m$, reducing the reference quality to price ratio. This renders the price of the good less salient and causes the consumer to focus on its high quality. This consumer's valuation is

---

sensitivity, $\sigma(x, y) > \sigma(x + \epsilon, y + \epsilon)$ for any $x, y, \epsilon > 0$. Ordering means that a larger price difference makes price more salient. Diminishing sensitivity means that a given price difference is less salient at a higher price level. These properties find considerable support in the literature on perception (see BGS 2012).

[15] In particular, because the reference $(q_t^m, p_t^m)$ is determined by the entire choice set, salience yields menu effects, such as the decoy effect, whereby the introduction of a new option can change the utility ranking of two pre-existing options. See BGS (2013) for details.

thus inflated relative to that of a consumer with a lower price norm. In the next section we explore in greater detail the patterns of choice that arise from (5).

Before moving on, we mention two reasons for choosing salience as a model of reference dependent choice, as opposed to loss aversion (Kahneman and Tversky 1979). First, Kahneman and Miller's idea of surprise does not feature an asymmetry between gains and losses: attributes can be surprising, and thus over-weighted in judgment, when they are far from their reference in either direction (and conversely, may not be surprising even if they are below the reference). Second, by shaping valuation through the perception channel, salience allows for truly irrelevant alternatives to affect choice. In contrast, loss aversion, at least in its original sense, can only be felt relative to past, expected, or aspired consumption.[16] For instance, in this approach choice cannot be influenced by past exposure to goods that were not chosen, whereas in our model such exposure shapes both norms and choices.

### 3.1. Buying Water Downtown and at the Airport

We illustrate the model in the simplest setting in which a consumer chooses between buying water of quality $q$ and the outside option $(0,0)$ of not buying it. Water costs $p$ downtown and $p+\Delta$ at the airport. At time $t$, the consumer faces the choice set $\boldsymbol{C}(t)=\{(q,p_t),(0,0)\}$, identified by the current price of water $p_t$. The set of norms at time $t$ is $\{(q,p_t^m),(0,0)\}$, so the reference is $(\bar{q}_t^m,\bar{p}_t^m)=\left(\frac{q}{2},\frac{p_t^m}{2}\right)$ and, from Equation (5), the consumer evaluates the option of buying water at time $t$ as:

$$\sigma\left(q,\frac{q}{2}\right)\cdot q-\sigma\left(p_t,\frac{p_t^m}{2}\right)\cdot p_t. \tag{6}$$

---

[16] Simonson and Tversky (1992) offer a model of the "background context" effect in which consumers are loss-averse relative to quality-price trade-offs observed in past choices. There are other models of selective attention where the weight of different attributes depends on the choice menu (Cunningham 2013, Koszegi and Szeidl 2013, Bushong, Rabin and Schwartzstein 2016). These models do not allow for a role of past choices to influence attention (with the exception of Cunningham 2013). A related phenomenon is coherent arbitrariness (Ariely, Loewenstein, Prelec 2003), in which experimental subjects' valuation for goods seems to be, to some extent, shaped by arbitrary anchors previously associated with those goods.

The model's predictions are then straightforward: price and quality are equally salient if both coincide with their normal levels (so both are proportionally equally distant from the reference levels). A good with normal quality $q$ is perceived to be a bad deal, and its price is salient, if it is abnormally expensive, $p_t > p_t^m$. The same good is perceived to be a good deal, and its quality is salient, if $p_t < p_t^m$.[17]

Consider the consumer who bought water only downtown $n$ times in the past. The price norm for water is $p_t^m = p$, irrespective of the current price. We then have:

**Proposition 3** *Given the set of norms* $\{(q,p),(0,0)\}$*, the consumer:*

*i) behaves rationally downtown, buying water if and only if* $q \geq p$.

*ii) overweighs price at the airport, and buys water if and only if* $q \geq \kappa_t^a \cdot (p+\Delta)$*, with* $\kappa_t^a > 1$.

Downtown, the price norm and the actual price of water coincide. The consumer is fully adapted. Price and quality are equally salient, and behavior is rational. At the airport, in contrast, the price of water is surprisingly high relative to the price norm. Price is salient and the consumer fails to buy even when a rational agent would. The low reference price acts as an irrelevant anchor that draws the consumer's attention to the current price. Thus, price is overweighed, distorting choice at the airport.

The prediction that the consumer is reluctant to buy water on the first airport visit is not unique to our memory based reference points. It could occur under mechanically adaptive reference points, but also under expectations-based reference points if the consumer has no prior exposure to, or knowledge about, airport prices (in which case the expected water price is also equal to $p$). Relative to these alternatives, however, memory based reference points have distinct predictions for the consumer's behavior *after* the first airport visit. This experience changes the consumer's memory database to the one in Equation (2).

---

[17] Formally, this results holds when $p_t > p_t^m/4$. The quality to price ratio logic obtains under the stronger condition $p_t > p_t^m/2$. These conditions are satisfied whenever the consumer is fully adapted (so that $p_t^m = p_t$), or provided the price norm is not much higher than the observed price, which holds throughout our analysis. When $p_t > p_t^m/2$, the available water has higher quality and price than the reference (which includes the option of not buying water). Thus, the good is a bad deal relative to the reference, $\frac{q}{p_t} < \frac{q/2}{p_t^m/2}$, when $p_t > p_t^m$, which implies that the good's price disadvantage is salient. Conversely, the good is a good deal when $p_t < p_t^m$, which implies that its quality is salient.

**Proposition 4.** *After the first airport visit at t, the consumer observes price* $p_{t'} \in \{p, p+\Delta\}$ *at* $t' > t$. *The price norm is then* $p_{t'}^{m} = p + w_t(p_{t'}) \cdot \Delta$, *where* $w_{t'}(p_{t'})$ *is the weight placed on the first airport experience. We then have:*

*i) observing* $p_{t'} = p + \Delta$ *yields a higher norm than observing* $p_{t'} = p$, *namely* $w_{t'}(p+\Delta) > w_{t'}(p) \geq 0$.

*ii) salience of price is lower at t' than at t, namely* $\sigma(p_{t'}, p_{t'}^{m}/2) < \sigma(p_{t'}, p/2)$, *for* $p_{t'} \in \{p, p+\Delta\}$

*The consumer overweighs quality downtown. He still overweighs price at the airport, but less than on the first visit. For* $p \in [k_{t'}^{d} \cdot q, k_{t'}^{a} \cdot q]$, $k_{t'}^{d} < 1 < k_{t'}^{a}$, *the consumer buys water downtown but not at the airport.*

Because of partial adaptation, memory based reference points create choice instability. After the first airport visit, the price norm for water – and thus its reference price – adjusts upward. As in Proposition 2, due to similarity-based recall, it adjusts more at the airport than downtown. The implications for choice are intuitive. Downtown, the higher price norm acts as a decoy. In comparison to that experience, downtown water seems a better deal: the reference quality price ratio drops, making the quality of downtown water salient. At the airport, the higher price norm reduces the salience of the high price, so water seems a better deal here too. In both locations, then, the valuation of water increases. But since adaptation is partial, price remains salient at the airport, and the consumer might still not buy even if normatively he should.

This mechanism of partial adaptation can provide a foundation for instances of "backward looking" reference points in the literature. Genesove and Mayer (2001) show that home sellers set asking prices that are tilted towards the purchase price they paid, which they suggest forms a reference price. Memory based norms would predict something similar: not only is the purchase price very available for recall, it also crowds out through interference prices recently observed for other houses (due to similarity along intrinsic characteristics). DellaVigna et al (2017) offer evidence of adaptation by recipients of unemployment insurance, who search more intensely for jobs around dates in which UI income predictably drops, yet eventually reduce search efforts. The authors suggest that UI recipients hold a reference point for consumption that averages consumption levels in the recent past. This evidence also lines up with our model, which predicts that a stable history generates a status quo norm (through both recency and similarity), which fails to fully adapt to a shock on impact, but eventually adapts when the shock persists.

Partial adaptation cannot be accounted for when reference points are given by rational expectations. Suppose that the consumer learns that airport prices are higher during his first visit. His norm for airport prices would then quickly converge to $p + \Delta$, and his reference point at the airport would become entirely independent of his own history. This consumer is no longer surprised at either location, and his price sensitivity should be the same in both locations, contrary to point ii). Likewise, mechanically adaptive reference points do not account for situation-specific adaptation (point i), because they predict that history (including the first airport visit) should influence the reference price equally at all locations. This model would predict that price sensitivity is always higher at the airport. In our model, in contrast, as consumers become better adapted to both downtown and airport prices, price norms converge to actual prices in both locations. When this occurs, choice behavior is stable in each location.

To see this logic, consider the long run behavior of a consumer who spends most of his time downtown but periodically visits the airport. To capture the relative infrequency of airport trips, suppose that the consumer visits the airport every $\tau_a$ periods, while he buys water downtown every $\tau_d$ periods, with $\tau_a > \tau_d$. We further assume that airport visits occur exactly in between two consecutive downtown shopping experiences. Adaptation to price in one location is interfered with by memories of the other location, which are stronger if experienced more recently. Thus, when shopping downtown, memories of airport prices are most available when the consumer just came back from the airport, with similarity $S(0, \mathit{\Delta}, \tau_d/2)$, while memories of downtown prices have maximum similarity $S(0,0,\tau_d)$. Conversely, when shopping at the airport, the most available downtown prices have similarity $S(0, \mathit{\Delta}, \tau_d/2)$, while memories of airport prices have maximum similarity $S(0,0,\tau_a) < S(0,0,\tau_d)$.

The following result shows that the level of adaptation depends on the strength of similarity along hedonic dimensions relative to that of contextual dimensions (i.e. time).

**Proposition 5.** *Suppose that $\eta \to \infty$ so that the norm is the experience most similar to the cue. Suppose further that $S(0,0,\tau_d) > S(0,\Delta,\tau_d/2)$, so that the price norm downtown is $p^m(p) = p$. Then quality and price are equally salient downtown, and valuation is proportional to $q - p$ regardless of recent experiences. Behavior at the airport can be in one of two regimes:*

*i) If airport visits are frequent enough, $S(0,0,\tau_a) > S(0,\Delta,\tau_d/2)$, the price norm fully adapts $p^m(p+\Delta) = p + \Delta$. Valuation is proportional to $q - p - \Delta$, with the same price sensitivity as downtown.*

*ii) If airport visits are infrequent, $S(0,0,\tau_a) < S(0,\Delta,\tau_d/2)$, the price norm does not adapt $p^m(p+\Delta) = p$. The consumer is always surprised by the airport price and is more price sensitive there.*

Proposition 5 highlights the conditions for full adaptation of reference points in the long run. If the consumer shops downtown frequently enough, he has a fully adjusted low reference price there – even if he just came back from the airport. In this case, valuation downtown is stable and independent of the last observed price. In this simple choice of whether to buy water, full adaptation means rational choice.

Full price adaptation to the airport requires that price similarity beats recency effects in recall, and thus only obtains when airport experiences are frequent enough, i.e. $\tau_a$ low enough. In this case, the reference price is $p + \Delta$ at the airport and $p$ downtown, so memory based reference points resemble expectations-based reference points. [18] Conversely, if airport visits are infrequent, then recency effects are strong enough that the downtown price enters the price norm at the airport. Even though the consumer has visited the airport many times, and is perfectly aware that airport prices are high, he is still surprised by them because downtown price repeatedly acts as an irrelevant anchor.

---

[18] Under full adaptation, the consumer behaves rationally in both locations in this setting. This strong result is due to two special assumptions: the choice is between buying water or nothing, and salience is homogeneous of degree zero. Under more general conditions, salience distortions affect even a fully adapted consumer. To see this, consider a choice between two goods, say a cheap wine and an expensive wine. The price difference between the two wines is constant, but their price level is higher at the restaurant than at the store. As we show in BGS (2013), diminishing sensitivity of salience implies that a fully adapted consumer finds the given price difference less salient at the high restaurant prices. As a result, he displays lower price sensitivity at the restaurant than at the store. This consumer deviates from rationality in that his focus on price, and thus his choice, is not consistent *across* situations. However, even in this case full adaptation implies strong choice consistency within situation: the price sensitivity of the consumer is the same at the store and at the restaurant regardless of the consumer's recent past.

Thaler (1985) illustrates the role of adaptation of norms in a choice setting. A beachgoer offers his companion to buy beer from a nearby establishment, and asks for his willingness to pay. In contrast to the predictions of the rational model, people state higher a willingness to pay for beer that comes from a nearby resort than for one that comes from a nearby store, even though the final consumption experience, beer on the beach, is exactly the same. Thaler suggests that the location (resort vs store) acts as a cue that brings to mind the past prices experienced in similar locations. In this setup, Proposition 5 suggests that, if the beachgoer rarely visits resorts, the frequently encountered store prices come to mind even when asked about the resort. His adaptation is only partial and he may refuse to buy beer at the resort price.[19] The more often the beachgoer visits resorts, the more his norm is shaped by resort prices, and the more he is willing to pay at the resort – while all along having a low price norm (and low willingness to pay) at the store.

In sum, our model produces stability in preferences (full adaptation) under much broader circumstances than mechanically adaptive models, but also identifies conditions in which consumers are systematically surprised by prices, despite being familiar with them (partial adaptation). In this sense, our predictions provide a middle-ground between those of mechanically adaptive and rational expectations based reference points. This middle-ground is a direct reflection of the fundamental properties of associative memory on which our model is based. First, and in contrast to both of the other approaches, reference points are generated ex post, by a spontaneous recall process. Second, recall is shaped by attribute similarity and by context (e.g., time) similarity. While the latter fosters adaptation to the recent past, the former fosters adaptation to the present. By incorporating fundamental mechanisms of memory, our model can shed light on diverse evidence that motivated several of the previous approaches.

## 4. An application to the housing rental market

Using data from the Panel Study of Income Dynamics (PSID) on U.S. households, Simonsohn and Loewenstein (2006) present two key findings. First, movers to a given U.S. city pay, on arrival, rent levels

---

[19] In section 4, we derive the properties of willingness to pay, and show it is increasing in the recalled price norm.

that are closer to those in the city of origin, when compared to otherwise identical households. The rent paid on arrival increases with rent levels in the city of origin, controlling for income, family size, and other observables. Second, households that subsequently move again within their destination city make rental choices that are no longer shaped by prices in the city of origin. SL argue, verbally, that their findings require a departure from rationality in which choice is anchored to recently experienced price levels.

The combination of memory based reference points and salience can account for these findings, and yields two additional predictions. First, city of origin price should exert a smaller effect on rental choice for movers who have previously lived in a city with housing prices similar to destination city prices (just like past airport visits help the consumer adapt to airport prices). Second, by diminishing sensitivity of salience, the influence of city of origin prices should be stronger for households moving to cheaper cities. This last prediction highlights a distinctive property of the salience model due to the logic of decoy effects. The expensive rents recalled from the city of origin act as decoys in the cheaper destination city, making even relatively expensive apartments in the latter look like a good deal. We next formalize this setting in our model (Section 4.1) and then we take the predictions to the data (Section 4.2).

### 4.1 Willingness to pay rent

The cleanest way to measure price salience effects would be to take two otherwise identical movers to the same city who have previously lived in different cities, and compare the rent they now pay for apartments of identical quality. This comparison is possible, despite quality being fixed, because prices faced by different households may differ due to market search or bargaining. Renters coming from expensive cities would have higher price norms, and be less price elastic. As a consequence, they would end up paying higher rents for the same quality, generating Simonsohn and Loewenstein's finding in a very stylized setting.

To formalize these ideas, we study the willingness to pay rent by a salient thinker with a memory-based reference rent for an apartment of *given* quality $q$. The idea is that the salient thinker receives offers drawn from the city's price distribution for apartment quality $q$ and accepts rental prices below his willingness to pay. By shaping willingness to pay, the memory-based reference rent shapes the average rent paid by the household for quality $q$ (which is the object of our empirical analysis).

One objection to this approach is that renters do not choose whether a certain price is acceptable for a given quality $q$. Rather, they face a choice set in which better apartments are more expensive, and they choose housing by trading off quality and price. To deal with this concern, we control in our regressions for several proxies for quality. To the extent that these proxies capture a large share of actual quality differences, our analysis can be viewed as approximating the ideal experiment of eliciting willingness to pay. It is possible to study a version of the model with a quality-price tradeoff, but that would complicate the analysis without producing predictions that are substantially different from those that we test.

Consider then the following setting. Apartments are of given known quality $q$. Faced with the choice between renting at price $p$ and not renting, $\boldsymbol{C} \equiv \{(q,p),(0,0)\}$, the consumer recalls the evoked set $\boldsymbol{C}_e = \{(q^m,p^m),(0,0)\}$, where $q^m, p^m$ are the memory-based reference quality and price. All housing has the same quality $q$, so that $q^m = q$. The salient thinker's $WTP$ for an apartment is then defined as:

$$WTP(q,p^m) = \sup_p \sigma\left(q,\frac{q}{2}\right)\cdot q - \sigma\left(p,\frac{p^m}{2}\right)\cdot p > 0. \tag{8}$$

Going forward, we assume the functional form:

$$\sigma(x,y) = e^{(1-\delta)\frac{|x-y|}{x+y}},$$

which yields convenient linear-regression expressions for the model's predictions. For a rational consumer ($\delta = 1$), $WTP(q,p^m) = q$, which is independent of past rental experience. For salient thinkers, willingness to pay is generically different from $q$, a result that should be unsurprising given the analysis in Section 3.

As we show in the Appendix, willingness to pay has two key properties. First, it increases in the reference price (i.e. $WTP(q,p^m)$ increases in $p^m$). This follows from the ordering property of salience, and the associated decoy logic of Proposition 3: a high reference $p^m$ acts as a decoy for the actual rent, rendering it less salient. This effect increases willingness to pay for $q$.[20] Second, willingness to pay is concave in the reference price (i.e., $WTP(q, p^m) - WTP(q, p^m - \Delta)$ is larger than $WTP(q, p^m + \Delta) - WTP(q, p^m)$). This effect is due to the diminishing sensitivity property of salience. A given price difference is more salient at lower price levels. Thus, the effect of city of origin prices on WTP should be stronger at lower price levels.[21]

To map the results on willingness to pay $WTP(q, p^m)$ to the data on movers, we write a mover's reference price $p^m$ in terms of the rent levels in his destination city, $p_d$, and in his city of origin, $p_o$,

$$p^m = p_o + w(p_d)(p_d - p_o), \tag{9}$$

where $w(p_d)$ is the weight that the norm puts on current prices $p_d$ relative to city of origin prices $p_o$. Note that we take the price cue to be the average price observed in the destination city and memory retrieval to occur with respect to average price levels observed in other cities in the past. This simplifies the model without altering its predictions on the behavior of the average mover. Our predictions for movers' behavior build on comparative statics of their rent norms, i.e. of $w(p_d)$, as a function of personal histories.

To derive testable implications, we log-linearize willingness to pay $WTP(q, p^m)$ by taking into account the Equation (9) for the normal price. We find the following result.

**Proposition 6.** *Under a log-linear approximation around the norm $p_0^m = q$, $WTP(q, p^m)$ satisfies:*

$$ln\, WTP = ln\, q \frac{\sigma}{\sigma + 2\sigma'} + \left(1 - w(p_d)\right) \frac{2\sigma'}{\sigma + 2\sigma'} ln\, p_o + w(p_d) \frac{2\sigma'}{\sigma + 2\sigma'} ln\, p_d, \tag{10}$$

*where $\sigma = \sigma(2,1)$, $\sigma' = \sigma'(2,1)$, and $\sigma' > 0$ if and only if $\delta < 1$.*

---

[20] This property holds for any $\sigma(x, y)$ provided $p^m$ is not much higher than $q$ (i.e, $p^m < q \cdot \frac{2\sigma(2,1)}{\sigma(1,1)}$), which ensures that $WTP > \frac{p^m}{2}$. The price component of utility, $\sigma\left(p, \frac{p^m}{2}\right) \cdot p$, is then monotonically decreasing in $p^m$ for $p$ close to $WTP$.

[21] Consider two households which recently experienced a rent level of $p^m = \$2000$ and which move to cities with rents levels of $\$1000$ and $\$3000$, respectively. The mover to the expensive city finds the higher price salient, but only moderately so. The mover to the cheap city, on the other hand, perceives a large price decline. The same $\$1000$ rental difference looms larger in the context of the cheap city price than in the context of the high expensive city price. This property requires the salience function not to be too concave, $2\sigma'(x, 1) + x \cdot \sigma''(x, 1) > 0$ for $x > 1$, which holds for the salience function above and for the salience functions considered in BGS (2012 and 2013).

By inspecting Equation (10), one can gauge the predictions of our model, which we bring to the data. All of these predictions obtain when holding constant the quality $q$ of the apartment.

*Prediction 1: Backward looking reference / Anchoring. On average, the rent paid after moving to the destination city increases in the rent level in the city of origin.*

For any $w(p_d) < 1$, WTP increases with rental levels in the city of origin $p_o$, as documented in SL (2006). Indeed, higher $p_o$ increases the reference price and willingness to pay.

*Prediction 2: Adaptation through recency. On average, if the household moves again in the destination city, the rent paid after this second move does not depend on city of origin price.*

Because the second time mover has experienced prices $p_d$ for a while, he is better adapted to the destination city, i.e. has a larger $w(p_d)$. Under full adaptation, $w(p_d) = 1$, the mover's rental expenditure no longer depends on the price of the city of origin ($p_o$ drops out of Equation (10)).

Predictions 1 and 2 were both tested in SL (2006). The following predictions are new.

*Prediction 3: Adaptation through similarity. Price in city of origin has a smaller effect on rent paid in the destination city for movers who had previously lived in cities with prices similar to $p_d$.*

Recall by price similarity causes movers previously exposed to price $p_d$ to be better adapted to the destination city than movers who have never experienced such a price. Formally, they have a larger $w(p_d)$. As a consequence, their expenditure is less anchored on $p_o$. Estimating Equation (10) for such movers thus yields a smaller coefficient on $ln\, p_o$ than for the full population of movers (i.e., in Prediction 1).

The last prediction of our model follows from concavity of $WTP(q, p^m)$. This prediction, which is proved in the Appendix, cannot be directly seen from the linearized Equation (10).

*Prediction 4: Asymmetry. Price in city of origin has a stronger effect on rent paid for movers to cheaper cities than for movers to more expensive cities.*

Formally, the coefficient on city of origin price (i.e., on $ln\, p_o$), should be higher for movers to cheaper cities than for movers to more expensive cities. While not a formal test of salience versus loss aversion (which would predict a strong reaction to price increases), this last prediction highlights a distinctive decoy effect property of the salience mechanism: raising the reference price makes high observed prices less salient, raising the good's valuation.

### 4.2 Empirical Tests

We use data from the Panel Study of Income Dynamics (PSID), a longitudinal yearly survey on a representative sample of U.S. families that also collects information on demographics and housing history over time. PSID data on housing history is now available from 1983 to 2013, roughly tripling the SL sample (1983-1993). [22] We supplement this data with historical data on median rents at the county level from the Fair Market Rents Dataset.[23] Like SL, we focus our analysis at the level of Metropolitan Statistical Areas (MSAs), so we use the terms city and MSA interchangeably. Median rents are aggregated to MSA level using population weights and all prices are converted to 1999 dollars.

We now describe the empirical strategy that we use to test predictions 1 to 4. Our analysis follows closely Simonsohn and Loewenstein's test of prediction 1. In implementing Equation (10), an observation is a household $i$ who moves in survey year $t$ and is a renter after the move. We use a household's post-move rent at year $t$, denoted $p_{it}$, as a proxy for their unobserved $WTP_{it}$.[24] We then run regressions of the form:

$$ln\, p_{it} = \beta_o \cdot ln\, p_{o,t_i} + \beta_d \cdot ln\, p_{d,t} + \boldsymbol{\beta_X} \cdot \boldsymbol{X}_{i,t} + \varepsilon_{i,t} \tag{11}$$

---

[22] The analysis uses data from PSID's Sensitive Data Files. We obtained access to this data under special contractual arrangements designed to protect the anonymity of respondents. PSID data is not available from the authors. PSID did not collect data on rent paid during the years 1988 and 1999, so these years are excluded from the analysis. We further trim the data in line with SL, and in particular focus on households observed for at least five survey waves and who move cities at least once. See Appendix B for details, and for differences in our approach and SL.

[23] Fair Market Rents data are available from the U.S. Department Housing and Urban Development (HUD), https://www.huduser.gov/portal/datasets/fmr.html.

[24] While $p_i$is a lower bound for $WTP_i$, this discrepancy should not systematically distort the predicted correlation with past prices (and conversely, it does not generate a spurious correlation with past prices in the rational benchmark).

Prices $p_{o,t_i}$ and $p_{d,t}$ denote the median rents in the household's city of origin and in the city of destination, respectively. Importantly, while rent levels in the current city are measured in the year of the move, $t$, rent levels in the city of origin are measured the last year the household lived there. Relative to Equation (10), the estimated parameters on rental prices correspond to $\beta_o = \left(1 - w(p_d)\right)\frac{2\sigma'}{\sigma+2\sigma'}$ and $\beta_d = w(p_d)\frac{2\sigma'}{\sigma+2\sigma'}$.

To estimate equation (11), we need to address two related econometric concerns. First, in our analysis apartment quality must be held constant. Second, we must address heterogeneity among households. Movers may be systematically different from stayers in several ways, including in their taste for housing, and these differences – not price experience per se – may be responsible for their different behavior. We address these concerns in the same way SL do. To control for housing quality, as well as for sources of household heterogeneity, we include in our regressions all standard variables that are used in regressions for housing demand and that are available in the PSID: household income, family composition, and age and education of head of household. We further account for unobserved differences across households by using information on households' previous choices. In particular, we control for whether the household previously rented or owned, as well as for a measure of relative taste for housing, namely the ratio $p_{i,t_i}/p_{o,t_i}$ of their rent expenditure to the median rent in the city of origin for past renters, and the analogous ratio in terms of house prices for past owners. Finally, we also include year fixed effects and a Heckman correction to account for the fact that, when households move, they endogenously select into renting, as opposed to buying. These controls help mitigate concerns about the selection of movers.

We test predictions 1 to 4 by estimating the regression (10) in the appropriate samples, which we now describe in detail. We test prediction 1 on backward looking reference points by using all observations of households in the year they move across cities. To test prediction 2 on adaptation through recency we consider households whom we observe moving within a city after having moved across cities. To test prediction 3 on adaptation through price-similarity, we focus on movers for whom we observe two moves across three cities. Because our prediction focuses on the second move, we refer to these cities as "earlier city", city of origin, and destination city. We measure price similarity between the earlier city and destination city by the absolute difference in median rent $|p_d - p_{earlier}|$. We then divide these movers into households for whom price similarity between destination and earlier cities is higher or lower than the median in this sample, and run the regression separately for each group. Finally, we test prediction 4 on asymmetry by dividing the baseline sample (used in Prediction 1) into those households who moved to more expensive versus cheaper cities.

Table I presents descriptive statistics of our samples' demographics, measured the year prior to their move. The samples are comparable in these dimensions. Households are equally likely to move "up" (to more expensive cities) as to move "down" (to cheaper cities), and face significant changes in rent levels ($152.6 on average, with $156.8 if moving up and $148.9 if moving down).

| | Head's Age (yrs) | Head's Education | Household Income ($) | Nr. Adults | Nr. Children | Median city rent ($) |
|---|---|---|---|---|---|---|
| Movers (N=2773) | 34.6<br>(14.3) | 14.1<br>(2.4) | 41,765<br>(37,117) | 1.64<br>(0.64) | 0.82<br>(1.19) | 652.38<br>(190.74) |
| Movers moving up (N=1,333) | 34.5<br>(13.2) | 14.15<br>(2.3) | 40,369<br>(32,225) | 1.61<br>(0.60) | 0.79<br>(1.14) | 570.34<br>(150.65) |
| Movers moving down (N=1,440) | 34.04<br>(12.67) | 14.09<br>(2.46) | 41,699<br>(31,646) | 1.64<br>(0.64) | 0.77<br>(1.15) | 739.30<br>(198.54) |
| Multiple Moves (N=504) | 33.81<br>(11.03) | 14.18<br>(2.27) | 41,101<br>(27,609) | 1.63<br>(0.61) | 0.91<br>(1.25) | 468.82<br>(338.73) |

Table I: Descriptive Statistics for Renters prior to move, at time $t-1$.

Table II presents the results. The estimates show the expected positive relation between rent paid and income, family size and local price levels. Intuitively, richer and larger households are likely to rent apartments of higher quality (e.g., larger ones). Focusing on the regressor of interest, $log(p_o)$, the results support predictions 1 and 2, and quantitatively confirm the results of SL (2006) in our larger dataset. In the baseline case (column 1), the coefficient $\beta_o$ on $log(p_o)$ is significantly positive and similar in magnitude to SL's: two otherwise identical individuals whose $p_0$ differs by one standard deviation differ in their rental expenditures in the same city by 3.4%. Prediction 2 also finds empirical support: when households move again within the same city (column 2), past city prices are no longer relevant. However, with the smaller sample size, we cannot conclude that this coefficient is significantly different from the baseline case.

|  | Backward looking reference | Adaptation through recency | Adaptation through price similarity |  | Asymmetry |  |
|---|---|---|---|---|---|---|
|  |  |  | Dissimilar | Similar | Moving up | Moving down |
| Log(income) | 0.253***<br>(0.0367) | 0.483***<br>(0.0346) | 0.339***<br>(0.0486) | 0.223***<br>(0.0590) | 0.416***<br>(0.0256) | 0.385***<br>(0.0229) |
| Nr. Children | 0.0475***<br>(0.0109) | 0.0566**<br>(0.0177) | 0.0518*<br>(0.0221) | 0.0815**<br>(0.0298) | 0.0511***<br>(0.0120) | 0.0481***<br>(0.0110) |
| Nr. Adults | 0.174***<br>(0.0240) | 0.152***<br>(0.0360) | 0.171***<br>(0.0375) | 0.187***<br>(0.0506) | 0.188***<br>(0.0254) | 0.167***<br>(0.0224) |
| $log(p_d)$ | 0.499***<br>(0.0499) | 0.583***<br>(0.0744) | 0.627***<br>(0.0983) | 0.589***<br>(0.137) | 0.524***<br>(0.0760) | 0.525***<br>(0.0783) |
| $log(p_o)$ | **0.163*****<br>(0.0458) | **0.0723**<br>(0.0557) | **0.221***<br>(0.106) | **0.173**<br>(0.141) | **0.0703**<br>(0.0797) | **0.243*****<br>(0.0744) |
| $p_{i,t-1}/p_0$ | 0.0560***<br>(0.0124) | 0.0607**<br>(0.0202) | 0.0300*<br>(0.0128) | 0.194*<br>(0.0684) | 0.0264**<br>(0.00989) | 0.0645***<br>(0.0101) |
| Constant | -2.094***<br>(0.365) | -2.798***<br>(0.558) | -3.114*<br>(0.877) | -0.807<br>(1.012) | -1.999***<br>(0.439) | -3.065***<br>(0.403) |
| N | 2773 | 719 | 257 | 247 | 1333 | 1440 |

Table II: Results from regression (12), estimated at MSA level. Not shown: age of head of household, (age squared)/100, female head, attended college, year fixed effects, inverse Mills ratio. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

To test prediction 3, we restrict the sample to households that move twice (columns 3 and 4). Consistent with our prediction, when movers have experienced price levels in the past that are similar to current ones (column 4), the influence of city of origin price $log(p_o)$ on rental expenditure in the destination city is insignificant. When movers have instead not experienced similar prices in the past (column 3), the effect of past prices is larger, and statistically significant. Again, given the small sample, the coefficients are not significantly different from each other.

Finally, in line with prediction 4, the anchoring of rents paid to past prices is driven almost entirely by households that move to cheaper cities, and rent more expensive apartments than locals do (columns 5 and 6). Past prices matter much less when otherwise similar households move to more expensive cities. The $\beta_o$ coefficients are different across the two samples at the 5% significance level.[25]

---

[25] The results of Table II are robust to different choices of specification (not shown). Controlling for endogenous selection into renting or for taste for housing, or excluding households who move for housing related reasons, plays essentially no role. Restricting the sample to households who rented *before* the move has little effect, except for prediction 3: while the results remain directionally consistent, the effect on households who experienced dissimilar prices is no longer significant, perhaps due to the much smaller sample size. Simonsohn and Loewenstein (2006) test a version of Prediction 4 to address concerns about learning, and find no asymmetry in their shorter sample.

The point estimates of Table II allow us to back out the weights in Equation (11). Using the fact that $\beta_o = \left(1 - w(p_d)\right)\frac{2\sigma'}{\sigma+2\sigma'}$, $\beta_d = w(p_d)\frac{2\sigma'}{\sigma+2\sigma'}$, and $\beta_o + \beta_d = \frac{2\sigma'}{\sigma+2\sigma'}$ we can back out, for each regression specification, the weight attached by the memory-based rent norm to city of destination price as $w(p_d) = \frac{\beta_d}{\beta_o+\beta_d}$. The baseline estimates of column 1 imply $w(p_d) = 0.754$, which is an average of the memory weight of different movers across different histories.[26] On the other hand, the point estimates from columns 2 and 4 of Table II allow us to assess the mechanisms for adaptation. These estimates respectively imply $w(p_d) = 0.890$ and $w(p_d) = 0.773$. Adaptation to current prices is stronger for movers who have spent time in the new city, or who have lived in similar cities in the past. The fact that the weight attached to city of destination price is higher for the former type of movers suggests that in our sample recency effects are stronger than price similarity effects. On the other hand, price similarity effects are sufficiently strong to generate significant patterns in the data. From column 3, movers who spent time in dissimilar cities are less adapted $w(p_d)_{diss} = 0.739 < w(p_d) = 0.890$.

To conclude, the evidence is consistent with the predictions of the model. Memory-based reference points provide a rationale for anchoring to recent rent levels (predictions 1 and 2), which were documented by SL, and also in Simonsohn (2006). Adaptation based on price similarity (Prediction 3) is a more nuanced prediction, and the evidence is statistically weaker but consistent with this prediction as well. Prediction 4 allows for a test of reference-dependent valuation, and again we find some support. The broader message is that our model generates novel predictions that can be tested using heterogeneous consumer experiences.

## 5. Conclusion

In this paper, we tried to make four contributions. First, we showed that one can incorporate a biologically founded, textbook model of memory (Kahana 2012) into an economic model of choice. The

---

[26] We could also try to estimate $\delta$ through the equality $\beta_o + \beta_d = \frac{2\sigma'}{\sigma+2\sigma'}$. However, the different average rental prices in different subsamples generates variation of salience $\frac{2\sigma'}{\sigma+2\sigma'}$ across these subsample, making it more difficult to back up parameter $\delta > 0$ without adjusting for these different price levels.

critical feature of this model – recall through similarity – yields many predictions on what comes to mind when decision makers face a stimulus, which have been extensively tested and confirmed in memory research but which also have multiple implications for economic analysis.

Second, we showed that this standard theory of recall naturally leads to a theory of memory based reference points. Due to the central role of similarity in recall, these reference points can often incorporate normatively irrelevant features, and through this channel lead to unstable and apparently irrational choice. But we showed that the same standard features of memory that explain irrelevant anchors lead to eventual adaptation of reference points that makes them situation-specific, and thereby creates the stability (and even rationality) of choice that is often observed. This approach to reference points can account both for some of the evidence on backward looking reference points, and some of the situations where reference points look like rational expectations.

Third, we combined the theory of memory based reference points with the salience theory of choice, which is a natural way to incorporate the notions of surprise, and over-reaction to surprise, into the theory of choice. Surprise relative to norms is critical to Kahneman and Miller's theory, and it emerges naturally from a combination of a textbook model of memory and salience theory.

Finally, we took the predictions to the data on movers between US cities, extending the work of Simonsohn and Loewenstein (2006). Our model predicts their basic findings, which we replicate with 20 additional years of data, but also yields additional predictions, for which we also find some support. Critically, these predictions come in part from our theory of choice, but also from the basic model of memory that we rely on throughout our analysis.

Throughout this paper, we have made a number of specific modeling choices for clarity, many of which can be revisited or relaxed. There are several missing aspects in the basic model of memory, such as the importance of salient memories, the inattention to some aspects of the initial stimulus that may influence recall (Schwartzstein 2014), or even the failure of initial encoding of some experiences. In addition, with some modifications, our model can perhaps also incorporate recall of other types of information from memory, such as goals or information about future events. In this sense, it may help to think about expectations as reference points, and in particular when expectations (as opposed to other information) are top of mind. In fact, we would argue that even the rational inattention approach (Sims 2003, Gabaix 2014) needs a theory of where the inputs into a decision not to pay attention come from, and recall of past conditions is likely to shape these inputs. We would also argue that phenomena involving the construction of preference such as projection bias, attribution bias, or the influence of past experiences on choice are all centrally related to memory. In this sense, portable textbook models of memory offer an opportunity to complete many different behavioral models and to improve their empirical testability.

## References

Ariely, Dan, George Loewenstein, and Drazen Prelec. 2003. "'Coherent Arbitrariness': Stable Demand Curves without Stable Preferences." *Quarterly Journal of Economics* 118 (1): 73 – 106.

Barberis, Nicholas and Ming Huang. 2001. "Mental Accounting, Loss Aversion, and Individual Stock Returns." *Journal of Finance* 56(4): 1247–1292.

Barberis, Nicholas and Wei Xiong. 2009. "What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation." *Journal of Finance* 64(2): 751–784.

Bell, David. 1985. "Disappointment in Decision Making Under Uncertainty." *Operations Research* 33(1): 1—27.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2012. "Salience Theory of Choice under Risk." *Quarterly Journal of Economics* 127 (3): 1243 -- 1285.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2013. "Salience and Consumer Choice." *Journal of Political Economy* 121(5): 803 -- 843.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." *Quarterly Journal of Economics* 131(4): 1753 -- 1794.

Brown, Gordon, Nick Chater, Ian Neath. 2007. "A Temporal Ratio Model of Memory" *Psychological Review* 114 (3): 539–576.

Bushong, Benjamin and Tristan Gagnon-Bartsch. 2016. "Learning with Misattribution of Reference Dependence." Unpublished Working Paper.

Bushong, Benjamin, Matthew Rabin and Joshua Schwartzstein. 2016. "A Model of Relative Thinking." Unpublished Working Paper.

Cunningham, Tom. 2013. "Comparisons and Choice" Unpublished Working Paper.

DellaVigna, Stefano, Attila Lindner, Balazs Reizer, and Johannes Schmieder. 2017. "Reference-Dependent Job Search: Evidence from Hungary." *Quarterly Journal of Economics*, forthcoming.

Dow, James. 1991. "Search Decisions with Limited Memory." *Review of Economic Studies* 58(1): 1-14.

Ericson, Keith and Andreas Fuster. 2014. "The Endowment Effect." *Annual Review of Economics* 6: 555-579.

Ericson, Keith. 2016. "On the Interaction of Memory and Procrastination: Implications for Reminders,Deadlines and Empirical Estimation." *Journal of the European Economic Association,* forthcoming.

Gabaix, Xavier. 2014. "A sparsity-based model of bounded rationality." *Quarterly Journal of Economics* 129 (4): 1661-1710.

Genesove, David and Christopher Mayer. 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market." *Quarterly Journal of Economics* 116(4): 1233 -- 1260.

Gennaioli, Nicola and Andrei Shleifer. 2010. "What Comes to Mind." *Quarterly Journal of Economics* 125 (4): 1399 -- 1433.

Gennaioli, Nicola, Andrei Shleifer and Robert Vishny. 2012. "Neglected Risks, Financial Innovation and Financial Fragility." *Journal of Financial Economics* 104 (3): 452 -- 468.

Gilboa, Itzhak and David Schmeidler. 1995. “Case-Based Decision Theory." *Quarterly Journal of Economics* 110(3): 605 -- 639.

Hastings, Justine and Jesse Shapiro. 2013. “Fungibility and Consumer Choice: Evidence from Commodity Price Shocks." *Quarterly Journal of Economics* 128(4): 1449 -- 1498.

Kahana, Michael. 2012. *Foundation of Human Memory*. Oxford University Press, Oxford UK.

Kahneman, Daniel and Amos Tversky. 1972. “Subjective Probability: A Judgment of Representativeness.” *Cognitive Psychology* 3 (3): 430–454.

Kahneman, Daniel and Amos Tversky. 1979. “Prospect Theory: an Analysis of Decision under Risk." *Econometrica* 47 (2): 263 -- 292.

Kahneman, Daniel and Dale Miller. 1986. “Norm Theory: Comparing Reality to its Alternatives.” *Psychological Review* 93(2): 136-153.

Kahneman, Daniel, Barbara Fredrickson, Charles Schreiber, and Donald Redelmeier. 1993. “When More Pain Is Preferred to Less: Adding a Better End. ” *Psychological Science* 4(6): 401-405.

Koszegi, Botond and Matthew Rabin. 2006. “A Model of Reference-Dependent Preferences", *Quarterly Journal of Economics*, 121 (4): 1133 -- 1165.

Koszegi, Botond and Adam Szeidl. 2012. “A Theory of Focusing in Economic Choice” *Quarterly Journal of Economics* 128 (1): 53-104.

Knutson, Brian, Scott Rick, Elliott Wimmer, Drazen Prelec, and George Loewenstein. 2007. “Neural Predictors of Purchases.” *Neuron* 53: 147–156.

Malmendier, Ulrike and Stefan Nagel. 2011. “Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?" *Quarterly Journal of Economics* 126(1): 373 -- 416.

Malmendier, Ulrike and Stefan Nagel. 2016. “Learning from Inflation Experiences." *Quarterly Journal of Economics* 131(1): 53-87.

Mas, Alexandre. 2006. “Pay, Reference Points, and Police Performance.” *Quarterly Journal of Economics* 121 (3): 783-821.

Mazumdar, Tridib, S. Raj, and Indrajit Sinha. 2005. “Reference Price Research: Review and Propositions.” *Journal of Marketing* 69 (4): 84 – 102.

Mullainathan, Sendhil. 2002. “Memory-Based Model of Bounded Rationality." *Quarterly Journal of Economics* 117 (3): 735-774.

Piccione, Michele, and Ariel Rubinstein. 1997. "On the interpretation of decision problems with imperfect recall." *Games and Economic Behavior* 20(1): 3-24.

Rubinstein, Ariel. 1998. *Modeling bounded rationality*. MIT press, Cambridge MA.

Schwartzstein, Joshua. 2014. “Selective Attention and Learning.” *Journal of the European Economic Association* 12 (6): 1423–1452.

Simonsohn, Uri and George Loewenstein, 2006. “Mistake #37: The Effect of Previously Encountered Prices on Current Housing Demand." *Economic Journal* (116)508: 175-199.

Simonsohn, Uri. 2006. "New Yorkers commute more everywhere: contrast effects in the field." *Review of Economics and Statistics* 88(1): 1-9.

Simonson, Itamar, and Amos Tversky. 1992. “Choice in Context: Tradeoff Contrast and Extremeness Aversion." *Journal of Marketing Research* 29 (3): 281-295.

Sims, Christofer. 2003. “Implications of Rational Inattention,” *Journal of Monetary Economics*, 50(3), 665–690.

Taubinsky, Dmitry. 2014. “From Intentions to Actions: A Model and Experimental Evidence of Inattentive Choice." Mimeo, Harvard University.

Thaler, Richard. 1985. “Mental Accounting and Consumer Choice”, *Marketing Science*, 4 (3): 199 – 214.

Torgerson, Warren. 1958. *Theory and Methods of Scaling*. John Wiley & Sons, Inc, New York, NY.

Tversky, Amos and Daniel Kahneman. 1973. “Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology* 5(2): 207-232.

Tversky, Amos. 1977. “Features of Similarity.” *Psychological Review* 84(4): 327-352.

Wilson, Andrea. 2014. "Bounded memory and biases in information processing." *Econometrica* 82(6): 2257-2294.

## Appendix A. Proofs

**Proposition 1.** Let $T_j$ be the set of time-stamps of all past experiences of good $j$ in the memory database $M_j$, and let $T_{(\hat{q},\hat{p})} = \{t: e_{jt} = (\hat{q},\hat{p},t)\} \subset T_j$ index past experiences of $(\hat{q},\hat{p})$. Then, the weight $w_j(\hat{q},\hat{p})$ is

$$w_j(\hat{q},\hat{p}) \; = \sum_{t\in T_{(\hat{q},\hat{p})}} \frac{h_{jt}}{\sum_{t_s\in T_j} h_{jt_s}}$$

The weights $h_{jt}$ assigned to individual experiences increase in their recency. It then follows from Property 2 of Definition 1 that $w_j(\hat{q},\hat{p})$ weakly increases if some or all experiences in $T_{(\hat{q},\hat{p})}$ become more recent. This proves point i.

Consider now changes in frequency of $(\hat{q},\hat{p})$ experiences, i.e. in the cardinality of $T_{(\hat{q},\hat{p})}$ (point ii). Suppose an additional experience $(\hat{q},\hat{p})$ occurs at $t'$, so that the new time-index set of such experiences becomes $T_{(\hat{q},\hat{p})} \cup \{t'\}$ . Then, the weight on $(\hat{q},\hat{p})$ increases provided

$$\frac{\sum_{T_{(\hat{q},\hat{p})}\cup\{t'\}} h_{jt}}{\sum_{T_{(\hat{q},\hat{p})}\cup\{t'\}} h_{jt} + \sum_{T_j/T_{(\hat{q},\hat{p})}} h_{jt}} > \frac{\sum_{T_{(\hat{q},\hat{p})}} h_{jt}}{\sum_{T_{(\hat{q},\hat{p})}} h_{jt} + \sum_{T_j/T_{(\hat{q},\hat{p})}} h_{jt}}.$$

This condition holds because $\sum_{T_{(\hat{q},\hat{p})}\cup\{t'\}} h_{jt} > \sum_{T_{(\hat{q},\hat{p})}} h_{jt}$ ■

**Proposition 2.** As in Proposition 1, the weight on experience $(\hat{q},\hat{p})$ given past experiences $T_{(\hat{q},\hat{p})}$ is given by

$$w_j(\hat{q},\hat{p}) = \sum_{t\in T_{(\hat{q},\hat{p})}} \frac{h_{jt}}{\sum_{t_s\in T_j} h_{jt_s}}$$

From Proposition 1, point ii, the weight $w_j(\hat{q},\hat{p})$ increases under the operation $T_{(\hat{q},\hat{p})} \to \hat{T}_{(\hat{q},\hat{p})} = T_{(\hat{q},\hat{p})} \cup \{\hat{t}\}$. Thus, $\widehat{w}_j(\hat{q},\hat{p}) > w_j(\hat{q},\hat{p})$, i.e. more weight is put on $(\hat{q},\hat{p})$ by the norm under memory database $\widehat{M}_j$ than by the norm under $M_j$.

Point ii states that the weight $\widehat{w}_j(\hat{q},\hat{p})$ increases with the similarity between $(\hat{q},\hat{p})$ and the cue $(q_t, p_t)$ is high. We have:

$$\widehat{w}_j(\hat{q},\hat{p}) = \frac{\sum_{t\in\hat{T}_{(\hat{q},\hat{p})}} h_{jt}}{\sum_{t\in\hat{T}_{(\hat{q},\hat{p})}} h_{jt} + \sum_{t_s\notin\hat{T}_{(\hat{q},\hat{p})}} h_{jt_s}}$$

which increases both in the frequency of exposure to $(\hat{q},\hat{p})$ (by Proposition 1, point i) and in the similarity of $(\hat{q},\hat{p})$ to $(q,p)$ (since the term $\sum_{t\in\hat{T}_{(\hat{q},\hat{p})}} h_{jt}$ increases while $\sum_{t_s\notin\hat{T}_{(\hat{q},\hat{p})}} h_{jt_s}$ stays fixed). ■

**Corollary 1.** Maximal interference under the ratio model leads to $w_{t_r} = 1$ if $t_r = argmax\{S(e_{t'}, e_{t_r})\}$, and $w_{t_r} = 0$ otherwise. To see this, recall that under the ratio model the weights equal $w_{t_r} = \frac{S(e_t,e_{t_r})^{\eta}}{\sum_{t_s\in T} S(e_t,e_{t_s})^{\eta}}$, and maximal interference is obtained for $\eta \to \infty$.

The assumption that similarity is stronger than recency, $S(0,0,|t'-t|) > S(|q_{t_r} - \hat{q}|, |p_{t_r} - \hat{p}|, |t_r - t|)$ for $t' < t_r < t$, implies that the most recent experience of $(\hat{q},\hat{p})$, at time $t'$, is more similar to the cue $(q,p)$ at time $t$ than other intervening experiences $(q_{t_r}, p_{t_r})$. Therefore, under maximal interference we have $w_{t'} = 1$. ■

**Proposition 3**. At time $t$, the reference price is $p^m = p$. Thus, the set of norms is $\{(q,p),(0,0)\}$ and the reference point, or average norm, is $\left(\frac{q}{2},\frac{p}{2}\right)$. The value of the outside option is 0. The value of water downtown is $\sigma\left(q,\frac{q}{2}\right)\cdot q - \sigma\left(p,\frac{p}{2}\right)\cdot p$ which, up to a constant, is equal to $q-p$. This proves point i.

At the airport, the value of water is $\sigma\left(q,\frac{q}{2}\right)\cdot q - \sigma\left(p+\Delta,\frac{p}{2}\right)\cdot(p+\Delta)$ so the traveler buys water if and only if

$$q > \frac{\sigma\left(1+\frac{\Delta}{p},\frac{1}{2}\right)}{\sigma\left(1,\frac{1}{2}\right)}\cdot(p+\Delta)$$

where $\frac{\sigma\left(1+\frac{\Delta}{p},\frac{1}{2}\right)}{\sigma\left(1,\frac{1}{2}\right)} > 1$ by the ordering property of salience (i.e. price is more salient than quality at the airport).

The result follows by setting $\kappa_t^a = \frac{\sigma\left(1+\frac{\Delta}{p},\frac{1}{2}\right)}{\sigma\left(1,\frac{1}{2}\right)}$, where $a$ stands for airport. ■

**Proposition 4**. Note first that, if $\eta < \infty$, then at time $t' > t$ the reference price $p^m(p_{t\prime}) \in (p, p+\Delta)$, for $p_{t\prime} \in \{p, p+\Delta\}$. In particular, we have $w_t(p_{t'}) > 0$ and $w_{t_n}(p_{t'}) > 0$ for either price realization (i.e. some positive weight is assigned to all past experiences). Moreover, $1 > w_t(p+\Delta) > w_t(p)$ because the recent airport price is more similar along the price dimension (and equally recent) to the current airport price, proving point i.

Because $w_t(p+\Delta) < 1$, the salience of price at the airport satisfies:

$$\sigma\left(p+\Delta,\frac{p+w_t(p+\Delta)\Delta}{2}\right) = \sigma\left(\frac{p+\Delta}{p+w_t(p+\Delta)\Delta},\frac{1}{2}\right) > \sigma\left(\frac{q}{q},\frac{1}{2}\right)$$

where the inequality follows from the ordering of the salience function. So price is still salient at the airport, though less so than at time $t$:

$$\sigma\left(\frac{p+\Delta}{p+w_t(p+\Delta)\Delta},\frac{1}{2}\right) < \sigma\left(\frac{p+\Delta}{p},\frac{1}{2}\right)$$

Downtown, we have that salience of price satisfies:

$$\sigma\left(p,\frac{p+w_t(p)\Delta}{2}\right) = \sigma\left(\frac{p}{p+w_t(p)\Delta},\frac{1}{2}\right) < \sigma\left(\frac{q}{q},\frac{1}{2}\right)$$

where the inequality holds – i.e. quality is salient – provided $\frac{p+w_t(p)\Delta}{2p} < 2$, which is guaranteed by our assumption that $p > \frac{p^m}{4} = \frac{p+w_t(p)\Delta}{4}$, see Footnote 17. This shows point ii.

Finally, setting $k_{t'}^{a} = \frac{\sigma\left(\frac{p+\Delta}{p+w_t(p+\Delta)\Delta},\frac{1}{2}\right)}{\sigma\left(\frac{q}{q},\frac{1}{2}\right)} > 1$ and $k_{t'}^{d} = \frac{\sigma\left(p,\frac{p+w_t(p)\Delta}{2}\right)}{\sigma\left(\frac{q}{q},\frac{1}{2}\right)} < 1$, the final result follows. ■

**Proposition 5**. The demonstration that reference points are fully adapted under the ratio model when $S(0,0,\tau_a) > S(0,\Delta,\tau_d/2)$ and $\eta \to \infty$ follows the steps of the proof of Proposition 2. Given that the consumer has fully adapted reference prices, i.e. $p^m(p_t) = p_t$ for $p_t \in \{p, p+\Delta\}$, it follows that $\sigma\left(q,\frac{q}{2}\right) = \sigma\left(p_t,\frac{p^m(p_t)}{2}\right)$ for $p_t \in \{p, p+\Delta\}$. As a consequence, valuation of the water $(q, p_t)$ is equal to $q - p_t$ (up to a normalization factor of $\sigma(2,1)$). This proves point ii. In particular, valuation is stable in that it does not depend on the recently observed prices, proving point i. ■

**Proposition 6.** We start by documenting some general properties of willingness to pay (WTP) in our model. Note that $WTP(q, p^m)$ is the largest solution $p$ to the following equation:

$$\sigma\left(q,\frac{q}{2}\right)q = \sigma\left(p,\frac{p^m}{2}\right)p$$

The right hand side is increasing in $p$ for $p > \frac{p^m}{2}$. As a consequence, a sufficient condition for the solution to this equation to be unique, for any salience function $\sigma$, is that $\frac{p^m}{2}$ is not too large, namely $\frac{p^m}{2} < q\frac{\sigma(2,1)}{\sigma(1,1)}$. We assume this condition going forward.[27] It then follows that $WTP(q, p^m) > \frac{p^m}{2}$.

First, $WTP(q, p^m) = q$ if and only if the reference price is $p^m = q$ (it follows from the above that this property is fully generic).

Second, $WTP(q, p^m)$ is increasing in $p^m$. Intuitively, for a given price $p > \frac{p^m}{2}$, the term $\sigma\left(p, \frac{p^m}{2}\right)p$ decreases in $p^m$, thus raising willingness to pay. Formally, define $\bar{q} = \frac{q}{2,}$ and $\bar{p} = \frac{p^m}{2,}$ , as well as $x \equiv p/\bar{p}$. Recall that, by assumption, $x > 1$. Then, we use homogeneity of degree zero, rewrite $WTP$ as the solution to the equation:

$$\bar{p}\sigma(x, 1)x = \hat{q},$$

where $\hat{q} \equiv \sigma(q, \bar{q})q$. From the implicit function theorem, the above defines a function $x(\bar{p})$ that satisfies

$$\frac{dx}{d\bar{p}} = -\frac{1}{\bar{p}}\frac{\sigma x}{\sigma + \sigma' x} < 0,$$

Where the inequality follows from the ordering property of salience, namely $\sigma' > 0$ for $x > 1$. This function also satisfies:

$$\frac{d^2x}{d\bar{p}^2} = -\frac{1}{\bar{p}}\frac{dx}{d\bar{p}}\left[2 - \frac{(2\sigma' + \sigma'' x)\sigma x}{(\sigma + \sigma' x)^2}\right]$$

We can now derive the comparative statics of reference price on the willingness to pay. Starting from $p = x\bar{p}$, we find:

$$\frac{dp}{d\bar{p}} = \frac{\sigma' x^2}{\sigma + \sigma' x} > 0,$$

---

[27] Uniqueness, and monotonicity, would hold for all $p^m$ under suitable conditions on $\sigma$ such that $\sigma\left(p, \frac{p^m}{2}\right)p$ is increasing in $p$.

so that, as advertised, willingness to pay increases in the reference price. Moreover, we find:

$$\frac{d^2p}{d\bar{p}^2} = x' \frac{(2\sigma' + \sigma'')\sigma x}{(\sigma + \sigma' x)^2}.$$

where $x' = \frac{dx}{d\bar{p}} < 0$. We thus find that WTP is concave in the reference price provided:

$$\frac{d^2p}{d\bar{p}^2} < 0 \Leftrightarrow 2\sigma' + \sigma'' > 0.$$

This condition states that the salience function should not be too concave. It is satisfied by our specification, $\sigma(x,1) = e^{(1-\delta)\frac{x-1}{x+1}}$ (where it reduces to $2\sigma' + \sigma'' \propto \frac{2x}{1+x} + (1-\delta) > 0$) as well as by the specifications we considered in previous papers, $\sigma(x,y) = \frac{x-1}{x+1}$.

We now derive a log-linear approximation to willingness to pay, around a reference price $p_0^m$ that satisfies $\bar{p} = \frac{p_0^m}{2} < q\frac{\sigma(2,1)}{\sigma(1,1)}$. Using the notation above, and taking logs we rewrite WTP as:

$$\ln p + \ln \sigma\left(\frac{p}{\bar{p}}, 1\right) = \ln q + \ln \sigma(2,1)$$

To first order in an expansion around a generic solution $\bar{p}_0, p_0$, the left hand side equals:

$$\ln p_0 + \left(\frac{p}{p_0} - 1\right) + \ln \sigma(x,1) + \frac{1}{\bar{p}_0}\frac{\sigma'(x,1)}{\sigma(x,1)}(p - p_0) - \frac{p_0}{\bar{p}_0^2}\frac{\sigma'(x,1)}{\sigma(x,1)}(\bar{p} - \bar{p}_0)$$

where $x = \frac{p_0}{\bar{p}_0}$. Because, by assumption, $\ln p_0 + \ln \sigma(x,1) = \ln q + \ln \sigma(2,1)$, these terms cancel out and the equation above becomes:

$$\left(\frac{p}{p_0} - 1\right) + x\frac{\sigma'(x,1)}{\sigma(x,1)}\left[\left(\frac{p}{p_0} - 1\right) - \left(\frac{\bar{p}}{\bar{p}_0} - 1\right)\right] = 0$$

Rewriting $\frac{p}{p_0} - 1 \approx \ln\frac{p}{p_0} = \ln p - \ln p_0$, and similarly $\frac{\bar{p}}{\bar{p}_0} - 1 \approx \ln p^m - \ln p_0^m$, replacing above, and regrouping, we find:

$$\ln p\,\big(\sigma(k,1) + x\sigma'(x,1)\big) = \ln p_0\,\sigma(x,1) + \ln p^m\, x\sigma'(x,1) + x\ln\frac{x}{2}\sigma'(x,1)$$

because $\frac{p_0}{p_0^m} = \frac{x}{2}$. Replacing again $\ln p_0$ for $\ln q - \ln \sigma(x,1) + \ln \sigma(2,1)$, the above becomes:

$$\ln p = \ln q \frac{\sigma(x,1)}{\sigma(x,1) + x\sigma'(x,1)} + \ln p^m \frac{x\sigma'(x,1)}{\sigma(x,1) + x\sigma'(x,1)} + \Lambda(x)$$

where

$$\Lambda(x) = x \ln \frac{x}{2} \sigma'(x,1) + \sigma(x,1)[\ln \sigma(2,1) - \ln \sigma(x,1)]$$

In Proposition 6 we considered the case $p^m = q = p$, which implies $\bar{p} = \frac{q}{2}$ and $x = 2$. In this case, $\Lambda(x) = 0$ and we get the expression in the text.

$$\ln p = \ln q \frac{\sigma(2,1)}{\sigma(2,1) + 2\sigma'(2,1)} + \ln p^m \frac{2\sigma'(2,1)}{\sigma(2,1) + 2\sigma'(2,1)}$$

Finally, note that the coefficient $\frac{x\sigma'(x,1)}{\sigma(x,1)+x\sigma'(x,1)}$ on the reference price is decreasing in $p^m$ if and only if the concavity condition derived above, $2\sigma' + \sigma'' x > 0$, is satisfied (as assumed above). ■